An Extended Nomenclature for Mammalian V-ATPase Subunit Genes and Splice Variants

Kevin C. Miranda¹,², Fiona E. Karet³, Dennis Brown¹,²*

¹Program in Membrane Biology and Division of Nephrology, Center for Systems Biology, Massachusetts General Hospital, Boston, Massachusetts, United States of America, ²Harvard Medical School, Boston, Massachusetts, United States of America, ³Department of Medical Genetics, University of Cambridge, Cambridge, United Kingdom

Abstract

The vacuolar-type H\(^{+}\)-ATPase (V-ATPase) is a multisubunit proton pump that is involved in both intra- and extracellular acidification processes throughout the body. Multiple homologs and splice variants of V-ATPase subunits are thought to explain its varied spatial and temporal expression pattern in different cell types. Recently subunit nomenclature was standardized with a total of 22 subunit variants identified. However this standardization did not accommodate the existence of splice variants and is therefore incomplete. Thus, we propose here an extension of subunit nomenclature along with a literature and sequence database scan for additional V-ATPase subunits. An additional 17 variants were pulled from a literature search while 4 uncharacterized potential subunit variants were found in sequence databases. These findings have been integrated with the current V-ATPase knowledge base to create a new V-ATPase subunit catalogue. It is envisioned this catalogue will form a new platform on which future studies into tissue- and organelle-specific V-ATPase expression, localization and function can be based.

Introduction

The vacuolar-type H\(^{+}\)-ATPase (V-ATPase) is a proton pump found in all nucleated cells of the body. Partially embedded within the membrane, its function is to transfer two hydrogen ions out of the cytoplasm at the expense of 1 ATP molecule [1]. It is thereby able to establish a proton gradient within the lumen of organelles such as lysosomes, endosomes and the trans-Golgi network. Organelle acidification is important for a diverse array of functions including intracellular trafficking and protein degradation [2–4].

The V-ATPase is also functionally important at the plasma membrane of specialized cell types in certain tissues. There, it is responsible for critical homeostatic functions such as body acid-base regulation (renal proximal tubule and collecting duct intercalated cells), bone remodeling (osteoclasts), and sperm storage and maturation (clear cells in the epididymis) [2,5,6] as well as other potential functions in other organs.

Loss of V-ATPase activity due to subunit mutations in specific cell types has been implicated in diverse pathophysiological states such as kidney and bone disease, sensorineural deafness and wrinkly skin syndrome [7–10]. Thus V-ATPase functions in a very broad array of physiological processes.

The V-ATPase is a large (800 kDa) and complex molecular motor. It is made up of at least 13 individual components/protein subunits organized into two functional domains: \(V_0\) and \(V_1\) [1,4,11,12]. The \(V_0\) domain is composed of several transmembrane subunits that are involved in hydrogen ion translocation across the bilayer, while the \(V_1\) domain is a peripheral to the membrane and is the site of ATP hydrolysis. \(V_0\) is composed of 5 subunits labeled a to e while \(V_1\) consists of 8 subunits denoted A to H. All 13 different subunits are encoded by separate genes located throughout the genome.

Many of the 13 V-ATPase subunits exist as homologs, thereby adding another level of complexity to the motor. The diverse functions and locations of V-ATPases are believed to be encoded within the various homologs. For example the d1 subunit is ubiquitously expressed, while the d2 homolog is seen only in the kidney, osteoclast and lung [13]. Similarly two isoforms of the B subunit were initially described as so-called “kidney” (B1) and “brain” (B2) specific isoforms, although it is now clear that expression of B1 is not restricted to the kidney [12]. Currently, homologs have been identified for the a, d, e, B, C, E and G subunits [11,14].

An additional level of V-ATPase subunit variation is encoded through splice variants. To date, splice variants have been identified for the a, d, e, C, G and H subunits [9,13–23]. Just like homologs, splice variants have been shown to exhibit different expression patterns. For example two splice variants of subunit a1 are expressed in rat neurons: one variant localized to axonal varicosities while the other was sorted to distal dendrites and axons [17].

To date, discovery of V-ATPase homologs and splice variants has largely been subunit focused and experimentally based. This method has proven successful as demonstrated by the large number of identified variants. However, this fragmented discovery process led to a fragmented naming system which was recently
Table 1. Overview of V-ATPase subunits and their respective RefSeq accession numbers.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Augmented Nomenclature</th>
<th>Nucleotide Accession</th>
<th>Ref</th>
<th>Sp</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP6V0E2</td>
<td></td>
<td>NM_001100592</td>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>ATP6V0E3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP6V0E20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The V-ATPase proton pump is composed of 13 subunits. According to the data presented in Smith et al 2003 these 13 are encoded through 22 homologs. These are denoted with a "K" for known in Table 1. A simple search of the literature and the Entrez Gene database [24] reveals another 17 variants. These are denoted with either an "E" for Entrez or "L" for literature in Table 1. All of these additional subunits are splice variants except for e2, which was cloned after the publication of Smith et al (2003).

A list of the current HUGO Gene Nomenclature Committee (HGNC) designations for the known V-ATPase subunits is provided, along with the corresponding RefSeq accession numbers.
another 4 potentially novel V-ATPase variants. Again, all new sequences are encoded by splice variants; these are denoted with a “D” for database search in Table 1. It should be noted that the additional 4 variants were cloned in high-throughput mRNA discovery studies. Thus, they have no accompanying experimental validation at the protein level; instead all exist as sequenced and apparently full length mRNA transcripts. Further experimental work will be required to verify the existence of these clones in vivo; nonetheless we have included them here. Thus, with this simple literature and database search we have increased the number of V-ATPase subunit variants from 22 to 43 (Figure 1). Finally, we propose that accessory proteins (whose functions related to the V-ATPase remain unknown) that also use the ATP6 nomenclature, including ATP6AP1 and ATP6AP2, should be included in the revised nomenclature scheme, and the appropriate suffixes should be added to their names if additional isoforms and splice variants emerge in the future. However, we have also examined these sequences and apart from pseudogenes, found no variants.

Experimentally identifying and characterizing the associated functional proteins was outside the scope of this study but should

**Figure 1. A schematic view of V-ATPase subunits.** The exonic structure of all V-ATPase subunits is shown; untranslated regions are shown in light grey while protein coding regions are black. The intronic distances have been divided by 15 for display purposes. Refer to Table 1 for further information on subunits.

doi:10.1371/journal.pone.0009531.g001
be the subject of future work by investigators interested in V-ATPase function, including our own groups. The results of this study suggest that the V-ATPase is regulated in a much more complex manner than is currently assumed. At what level this regulation is exerted remains to be determined experimentally, and having a complete systems overview of all V-ATPase components will help expedite this process. This almost two-fold increase in the number of V-ATPase subunit forms demonstrates how the output of broad genomic scale projects can be utilized for a specialized pursuit. It also highlights the importance of computational methods for sifting and sorting through vast amounts of data deposited in sequence databases worldwide. Identification and standardization of transcript variation offers a powerful approach to guide the future assignation of functional significance among protein variants.

Materials and Methods

Sequences corresponding to the 13 human V-ATPase subunits described previously [11] were retrieved from the HUGO Gene Nomenclature Database (HGNC) database [25]. These sequences were used to query the RefSeq non-redundant protein and nucleotide databases [24] with the BLAST algorithm [26]. Various modules from the Bioperl toolkit were utilized to process the resulting BLAST output [27]. The HGNC V-ATPase identifiers were used to query the SpliceCenter databases [28]. The results of this analysis were manually integrated with the BLAST results described above. This integrated catalogue is presented in Table 1 and Figure 1.

Author Contributions

Conceived and designed the experiments: KM. Performed the experiments: KM. Analyzed the data: KM FK DB. Contributed reagents/materials/analysis tools: KM FK DB. Wrote the paper: KM FK DB.

References


